All-photon Polarimetric Time-of-Flight Imaging

-Supplemental Document-

Seung-Hwan Baek[1,2] Felix Heide[1]

Princeton University[1] POSTECH[2]

In this Supplemental Document, we present additional details and analysis in support of the findings we reported in the main document. Specifically, we provide here

- additional details on the temporal-polarimetric reflectance and image formation,
- additional details on ellipsometric reconstruction,
- additional details on scene reconstruction,
- calibration discussion,
- an imaging system part list,
- additional details on temporal-polarimetric BRDF model,
- additional discussion on acquisition time and high-order light transport, and
- additional synthetic and experimental results.

1. Temporal-Polarimetric Reflectance Model and Image Formation

In this section, we provide an extended formulation of our reflectance model and the image formation which was presented in the main paper in concise forms.

1.1. Mueller Matrices of Optical Elements, Surface Reflection, and Coordinate Conversion

Our BRDF and image formation models rely on the Mueller matrices of various optical elements and light-matter interaction. We provide the individual matrices in this Supplemental Document for completeness. For further reading, we refer to Collet [4].

Linear Polarizer A linear polarizer placed at an angle θ with respect to a reference axis has the Mueller-matrix form

$$\mathbf{L} = \frac{1}{2}\begin{bmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \cos 2\theta \sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta \sin 2\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{1}$$

Wave Plate The Mueller matrix of a wave plate with retardance ϕ at angle θ to the horizontal is defined as

$$\begin{aligned} \mathbf{R} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathbf{R}_{11} & \mathbf{R}_{12} & \mathbf{R}_{13} \\ 0 & \mathbf{R}_{21} & \mathbf{R}_{22} & \mathbf{R}_{23} \\ 0 & \mathbf{R}_{31} & \mathbf{R}_{32} & \mathbf{R}_{33} \end{bmatrix}, \\ \mathbf{R}_{11} &= \cos^2 2\theta + \sin^2 2\theta \cos\phi, \\ \mathbf{R}_{12} &= \sin 2\theta \cos 2\theta (1 - \cos\phi), \\ \mathbf{R}_{13} &= \sin 2\theta \sin\phi, \\ \mathbf{R}_{21} &= \sin 2\theta \cos 2\theta (1 - \cos\phi), \\ \mathbf{R}_{22} &= \cos^2 2\theta \cos\phi + \sin^2 2\theta, \\ \mathbf{R}_{23} &= -\cos 2\theta \sin\phi, \\ \mathbf{R}_{31} &= -\sin 2\theta \sin\phi, \\ \mathbf{R}_{32} &= \cos 2\theta \sin\phi, \\ \mathbf{R}_{33} &= \cos\phi. \end{aligned} \tag{2}$$

Note that the ideal half/quarter-wave plates have a retardance values of π and $\pi/2$, respectively, resulting in the Mueller matrices

$$\begin{aligned} \mathbf{W}(\theta) &= \mathbf{R}(\theta, \phi \leftarrow \pi), \\ \mathbf{Q}(\theta) &= \mathbf{R}(\theta, \phi \leftarrow \pi/2). \end{aligned} \tag{3}$$

We use the calibrated retardance values of our QWP and HWP at the laser wavelength of 635 nm.

Non-polarizing Beam Splitter Even though not described in the main paper, our non-polarizing beam splitter also affects the polarization state of light. The beam splitter

has a minimal impact on the polarization of the transmitted light, resulting in the identity Mueller matrix

$$\mathbf{B}_{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

For the reflected light, the beam splitter negates the last two elements of the Stokes vector in the same fashion as an ideal mirror, that is

$$\mathbf{B}_{\mathrm{R}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{5}$$

Surface Fresnel Reflection and Transmission Light interacting with a smooth surface changes its polarization state according to Fresnel law. Specifically, the transmitted and reflected components are represented as the Fresnel Mueller matrices

$$\mathbf{F}_{T,R} = \begin{bmatrix} \frac{\mathrm{F}^{\perp}+\mathrm{F}^{\parallel}}{2} & \frac{\mathrm{F}^{\perp}-\mathrm{F}^{\parallel}}{2} & 0 & 0 \\ \frac{\mathrm{F}^{\perp}-\mathrm{F}^{\parallel}}{2} & \frac{\mathrm{F}^{\perp}+\mathrm{F}^{\parallel}}{2} & 0 & 0 \\ 0 & 0 & \sqrt{\mathrm{F}^{\perp}\mathrm{F}^{\parallel}}\cos\delta & \sqrt{\mathrm{F}^{\perp}\mathrm{F}^{\parallel}}\sin\delta \\ 0 & 0 & -\sqrt{\mathrm{F}^{\perp}\mathrm{F}^{\parallel}}\sin\delta & \sqrt{\mathrm{F}^{\perp}\mathrm{F}^{\parallel}}\cos\delta \end{bmatrix}. \tag{6}$$

We can compute the perpendicular and parallel Fresnel coefficients $F^{\perp,\parallel}$ for reflection and transmission [4], respectively. Here, δ is the phase shift that has the value of ϕ or 0 for the dielectric component. Specifically, if the incident angle is smaller than the Brewster angle, δ is ϕ, otherwise 0. The Fresnel Muller matrices used by us for surface and sub-surface reflections can be computed from this general Fresnel Mueller matrix by using the corresponding parameters of the refractive index of a material η, incident angle, and outgoing angle. A similar derivation can be found in Baek et al. [2].

Coordinate Conversion The Stokes-Mueller formalism is defined based on the coordinate systems used. Specifically, a Stokes vector is defined with respect to a coordinate. A Mueller matrix is defined for a pair of input Stokes coordinate and an output Stokes coordinate. This requires careful coordinate conversion for every operation applied to Stokes vectors and Muller matrices. A coordinate-conversion Mueller matrix can be used for this purpose and has the following form

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta & 0 \\ 0 & -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7}$$

where θ is the rotation angle of the input x basis vector. That is, multiplying the coordinate conversion matrix rotates the coordinate frame of the xy axes by the angle θ while the propagation direction of the z axis is maintained.

1.2. Modeling Intensity-Attenuation based on Micro-facet Distribution

We model surface reflection based on microfacet theory. Specifically, we use the Smith Shading and masking term G from [7] and the GGX facet distribution term D from [9] to represent realistic BRDFs. They are defined as follows

$$D(\theta_h;\sigma) = \frac{\sigma^2}{\pi\cos^4\theta_h(\sigma^2+\tan^2\theta_h)^2},$$
$$G(\theta_i,\theta_o;\sigma) = \frac{2}{1+\sqrt{1+\sigma^2\tan^2\theta_i}}\frac{2}{1+\sqrt{1+\sigma^2\tan^2\theta_o}}, \tag{8}$$

where θ_h is the half-way angle, θ_i and θ_o are the incident and outgoing angles. Here, σ is the surface roughness.

1.3. Complete Image Formation Model

In the main manuscript, we listed Muller matrices of the analyzing and polarizing optics. We provide here a non-factored formulation including the beam splitter and the galvo mirror, that is

$$\mathbf{A} = \mathbf{L}\mathbf{Q}\mathbf{B}_{\mathrm{R}}\mathbf{G},\ \mathbf{P} = \mathbf{G}\mathbf{B}_{\mathrm{T}}\mathbf{Q}\mathbf{W}, \tag{9}$$

where G is the Mueller matrix of the galvo mirror which can be modeled with Fresnel Muller matrices [8]. We plug the complete Mueller matrices of the analyzing and the polarizing optics into the original image formation model (Equation (6) in the main manuscript).

2. Additional Details on Ellipsometric Reconstruction

We perform least-square optimization to estimate the Mueller matrix of each temporal-spatial pixel from the polarimetric measurements with varying rotation angles, following Baek et al. [3]. That is, we first rearrange the captured intensities $I_{i\in\{1,\cdots,N\}}$ into a vector $\mathbf{I}$. We can then reconstruct the target Mueller matrix $\mathbf{H}$ using a pseudo-inverse as

$$\dot{\mathbf{H}} = (\mathbf{K}^{\intercal}\mathbf{K})^{-1}\mathbf{K}^{\intercal}\mathbf{I}, \tag{10}$$

where $\dot{\mathbf{H}}$ is the vectorized version of $\mathbf{M}$, and $\mathbf{K}$ is the stacked system Matrix for all rotation angles $\mathbf{K}_{i,\forall} = [\mathbf{A}_i]_{0,\forall}[\mathbf{P}_i]_{\forall,0}$.

3. Additional Details on Scene Reconstruction

Once we obtain the Mueller matrix $\mathbf{H}$ via the ellipsometric reconstruction, we now turn to reconstruct scene parameters by fitting our temporal-polarimetric rendering model

to the Mueller matrix $\mathbf{H}$. In the main paper, we formulated this as an optimization problem

$$\underset{\dot{\mathbf{\Theta}}}{\text{minimize}} \left\| \dot{\mathbf{W}}_p \odot \left(f(\dot{\mathbf{\Theta}}) - \dot{\mathbf{H}}_{\text{meas}} \right) \right\|_1 + \left\| \dot{\mathbf{W}}_d \odot \nabla \dot{\mathbf{n}} \right\|_1. \tag{11}$$

Reparameterizing Constrained Scene Parameters One of the key details for solving this optimization problem is to constrain the range of each scene parameter based on its physical interpretation and target scene configurations. While constrained optimization has been extensively studied in optimization, we present a simple reparameterization technique that we can incorporate into existing gradient-descent optimization frameworks. Specifically, for each scene parameter $\mathbf{x} \in \mathbf{\Theta}$, we apply a scaled logit function to obtain the corresponding optimization variable $\mathbf{x}_o$ as

$$\mathbf{x}_o \leftarrow \text{logit}\left(\frac{\mathbf{x} - v_{\min}}{v_{\max} - v_{\min}} \right), \tag{12}$$

where $v_{\min}$ and $v_{\max}$ are the minimum and maximum values manually set for the scene parameter $\mathbf{x}$. This operation is differentiable for the expected input range: $v_{\min} < \mathbf{x} < v_{\max}$, enabling gradient flow throughout the optimization process. We operate the Adam optimizer on this reparameterized optimization variable $\mathbf{x}_o$ instead of the scene parameter $\mathbf{x}$, allowing to constrain the range of the scene parameter. We convert the optimization variable $\mathbf{x}_o$ to the original scene parameter $\mathbf{x}$ when we predict the temporal-polarimetric Mueller matrix using the image formation model $f()$. We perform this inverse operation as

$$\mathbf{x} \leftarrow (v_{\max} - v_{\min}) \text{sigmoid}\left(\mathbf{x}_o\right) + v_{\min}. \tag{13}$$

We set the minimum and maximum values ($v_{\min}$ and $v_{\max}$) of the entire scene parameters as: $v_{\min} = 0.5, v_{\max} = 2$ for the depth in meters, $v_{\min} = 1.2, v_{\max} = 1.9$ for the refractive index, $v_{\min} = 0.1, v_{\max} = 2$ for the roughness, $v_{\min} = 0, v_{\max} = 0.8$ for the scattering mean μ, $v_{\min} = 0.004, v_{\max} = 0.4$ for the scattering standard deviation σ, and $v_{\min} = 0.01, v_{\max} = 4$ for the scattering scalar a.

We apply a unit-norm constraint to the surface normals. To this end, we normalize the unconstrained optimization variable of the surface normals when we use the rendering function f.

Weighted Optimization We introduce the weighting matrix $\mathbf{W}_d$ for preventing overly-smoothed surface normals around depth edges. Specifically, we compute the sum of x and y gradients of the depth map $\mathbf{d}$ and compute the confidence map for the smoothness term as

$$\mathbf{W}_d(x) = 0.0001(1 - \text{sigmoid}(10(|g_x(p)| + |g_y(p)| - 0.5))), \tag{14}$$

where x is a pixel, and g_x and g_y are the x and y gradients of the depth map $\mathbf{d}$. This weighting function evaluates to one if we do not evaluate on a depth edge but decreases as we get closer to the edges. The other weighting matrix $\mathbf{W}_p$ sets different weighting values for the Mueller matrix elements, which is useful as their scales are significantly different, especially for the diagonal and the non-diagonal Mueller-matrix elements. This is particularly important for real-world measurements where SNRs are different for each matrix element. We use 0.1 for the weighting values of all matrix elements in the synthetic evaluation. For the real experiments, we apply the weighting values of 0.3 for the first element $\mathbf{H}_{00}$ and 0.1 for the elements $\mathbf{H}_{22,33,10,20}$, and set the others to zero.

Material Clustering In our synthetic experiments, we extract the cluster maps directly from the Blender renderer. For the real-world scenes, given a manually-selected number of clusters, we perform k-means clustering on the vectorized steady-state intensity image $\sum_t \mathbf{H}_{00}(t)$.

4. Calibration

Locating accurate 3D scene points in a scene requires precise geometric calibration of the imaging system. We capture checkerboards at different positions and estimate the intrinsic parameters [11]. This allows for converting a depth map $\mathbf{d}$ into a point cloud $\mathbf{p}$ which we can use for computing the per-point incident and outgoing vectors $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_o$. We also perform polarimetric calibration to compensate for the mismatch between the image formation model and the real measurements due to non-ideal optics. Specifically, we capture an unprotected gold mirror (Thorlabs PF10-03-M03) with known Mueller matrix, i.e, an ideal mirror, and estimate the correction Mueller matrix [5].

5. Imaging System Parts List

We list all parts used to build the experimental prototype system in Table 1.

6. Details on Temporal-Polarimetric BRDF

Coordinate Conversion for Sub-surface Scattering For sub-surface scattering, we do a coordinate conversion of Stokes vectors in our temporal-polarimetric BRDF model. Specifically, for incident/outgoing light to/from a surface, we define the corresponding Stokes vector in a coordinate that has a z axis matching the light propagation direction, while the x and y axes can be arbitrarily chosen to form an orthonormal basis [4]. Among infinite candidates for the x and y axes, we choose them following the halfway coordinate approach [3]. In halfway coordinates, we define the y axis on a plane spanning the halfway vector $\mathbf{h}$ and the propagation direction, which determines x axis as $y \times z$. In this

Item #	Part description	Quantity	Model name
1	Singe-photon avalanche diode	1	MPD
2	Time-correlated single photon counting	1	PicoQuant TImeHarp 260 PICO
3	Objective lens	1	Canon EF 50mm f/1.8 STM
4	Linear polarizer	1	Newport 10LP-VIS-B
5	Achromatic quarter-wave retarder	2	Thorlabs AQWP10M-580
5	Achromatic half-wave retarder	1	Thorlabs AHWP10M-600
7	Motorized rotary stage	4	Thorlabs KPRM1E
8	Non-polarizing beamsplitter	1	Thorlabs BS013
9	Beamsplitter cage	1	Thorlabs CCM1-4ER
10	Dielectric mirror	1	Thorlabs BB1-E02
11	Right-angle mirror mount	1	Thorlabs KCB1
12	Beam block	1	Thorlabs LB1
13	Cage aperture	1	Thorlabs CP20S
14	Picosecond pulsed diode laser	1	Edinburgh Instruments EPL-635
15	Galvo mirror	1	Thorlabs GVS012
16	Galvo power supply	1	GPS011-US
17	EOS-to-C-mount adaptor	1	Fotodiox pro EOS (α)-C-D-Click

Table 1. Part list of our imaging system.

convention, we convert the halfway coordinate of the incident/outgoing light direction also to a surface normal coordinate. This results in the y axis to fall in a plane that spans the surface normal and the propagation direction because the transmission event itself is defined on the corresponding surface normal coordinate, see also [2].

Experimental Validation Validating the representation capability of a parametric BRDF model can be done by comparing the real BRDF measurements of many materials and the BRDFs predicted from the parametric model. However, this mandates the availability of sufficiently many real-world BRDF measurements, which is unfortunately not the case for temporal-polarimetric BRDFs. Instead, we conducted an experiment that validates the agreement of the real-world temporal-polarimetric measurements and our model prediction for a single spherical object. Figure 5 in the main manuscript shows the two temporal slices of the polarimetric measurements and our prediction with manually chosen parameter sets. Our temporal-polarimetric BRDF model allows us to match the two important aspects of the real-world BRDF: temporal-polarimetric changes of (1) specular highlight arising from surface reflection and (2) normal-dependent periphery due to sub-surface reflection. We note that using a perfectly *calibrated* material in terms of geometry, roughness, refractive index, and scattering parameters could be an alternative way for the BRDF evaluation even though it is also challenging to fabricate and calibrate such a material.

7. Additional Discussion

Improving Acquisition Time For our experiments, we use a low-speed galvo mirror (Thorlabs GVS012) and mechanically rotating polarizers (Thorlabs KRPM1e) available to us. As such, the acquisition speed may be drastically increased by using existing mega-pixel SPAD sensors and electronically-controllable liquid-crystal tunable filters.

High-order Light Transport Higher-order light transport including interreflection may degrade our reconstruction quality. We do not consider these transport components in our work as the proposed coaxial capture configuration suppresses the capture of these components.

8. Additional Results

8.1. Synthetic Experiments

For the synthetic experiments, we designed scenes in Blender similar to the popular *Cornell Box*. Specifically, we used models from the *Berlin Egyptian Museum*, *Stanford Bunny*, and *Stanford XYZ Dragon* [6]. We render depth, surface normals, and index maps and manually assign the material parameters of our reflectance model to each material index. In the main manuscript, we provided results on a synthetic dataset for the scene *Nefertiti*. We here present additional results for *Nefertiti*, *Bunny*, and *Dragon* scenes, shown in Figures 3, 4, and 5. Note that material properties and local surface curvature affect the reconstruction accuracy of material and geometry. We compute per-object (box and target) material reconstruction accuracy in Figures 3, 4,

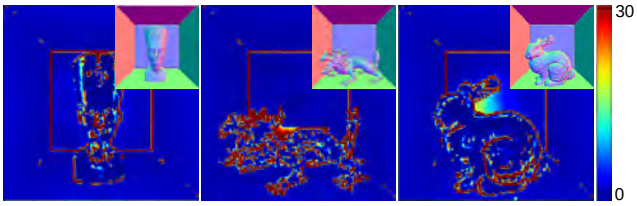


Figure 1. Error maps of the estimated normals for the three synthetic scenes in degrees. Our method recovers precise geometry except for occlusion edges and geometric discontinuities.

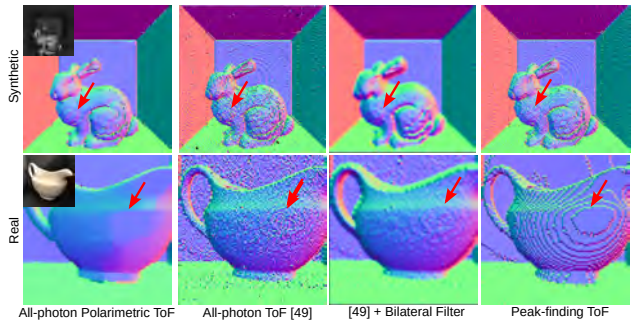


Figure 2. All-photon polarimetric imaging captures geometric discontinuities (red arrow) that cannot be resolved in non-polarimetric all-photon imaging [49], and also after bilateral filtering, both in simulation and experimentally.

and 5. In addition, we show the error maps for normal reconstruction in Figure 1 for the three scenes. Our method recovers accurate geometry for most regions except occlusion edges and discontinuities.

8.2. Real Experiments

We present the additional results on four real-world scenes (*Pot*, *Basketball*, *Dragon*, and *Candle*) in Figures 6, 7, 8, and 9. Specifically, we compare the estimated geometry from conventional peak-finding ToF, shape-from-polarization (SfP), and our polarimetric ToF imaging approach. The results validate the effectiveness of our all-photon polarimetric ToF imaging, which is capable of obtaining high-fidelity continuous geometry. We also visualize the time-varying angle-of-linear polarization (AoLP) and degree of linear polarization (DoLP) [3]. This visualization highlights the temporal-polarimetric structure of surface and sub-surface reflections. In additiona, we show the measured, initialized, and disentangled time-varying intensities. We find that the proposed polarimetric ToF method effectively decomposes the scene reflection into surface and sub-surface reflections in the temporal and polarimetric dimensions as also evidenced by the optimized temporal-polarimetric matrices $\mathbf{H}$. We further evaluate the experimental geometric precision of the proposed method by comparing our estimated geometry to ground truth geometry obtained with a structured-light scanning in the recent polarimetric probing method [1]. Specifically, we used the dragon statue shown in Figure 8. The average errors of the estimated surface normals are 45 and 23 degrees for the conventional peak-finding ToF and our method.

Comparison to Non-Polarimetric All-photon Imaging Non-polarimetric all-photon ToF imaging [10] fits a parametric temporal BRDF model to temporal measurements in order to estimate scene depth. Exploiting all temporal photons in the fitting process improves reconstruction accuracy of scene depth especially in the presence of measurement noise, compared to conventional peak-finding ToF imaging methods. However, its depth resolution is still fundamentally limited by the temporal resolution, which is generally low due to the hardware limitation of existing affordable temporal-sensing devices such as MPD SPAD. Our polarimetric all-photon ToF imaging tackles this problem by additionally incorporating the high-resolution geometry cue from polarization. It allows us to reconstruct geometric details from polarization as shown in Figure 2. We recover the vase crease, i.e., the geometric discontinuity indicated by the red arrow.

References

[1] Seung-Hwan Baek and Felix Heide. Polarimetric spatio-temporal light transport probing. *ACM Transactions on Graphics (TOG)*, 2021. 5

[2] Seung-Hwan Baek, Daniel S Jeon, Xin Tong, and Min H Kim. Simultaneous acquisition of polarimetric svbrdf and normals. *ACM Transactions on Graphics (TOG)*, 37(6):268–1, 2018. 2, 4

[3] Seung-Hwan Baek, Tizian Zeltner, Hyunjin Ku, Inseung Hwang, Xin Tong, Wenzel Jakob, and Min H Kim. Image-based acquisition and modeling of polarimetric reflectance. *ACM Transactions on Graphics (TOG)*, 39(4):139, 2020. 2, 3, 5

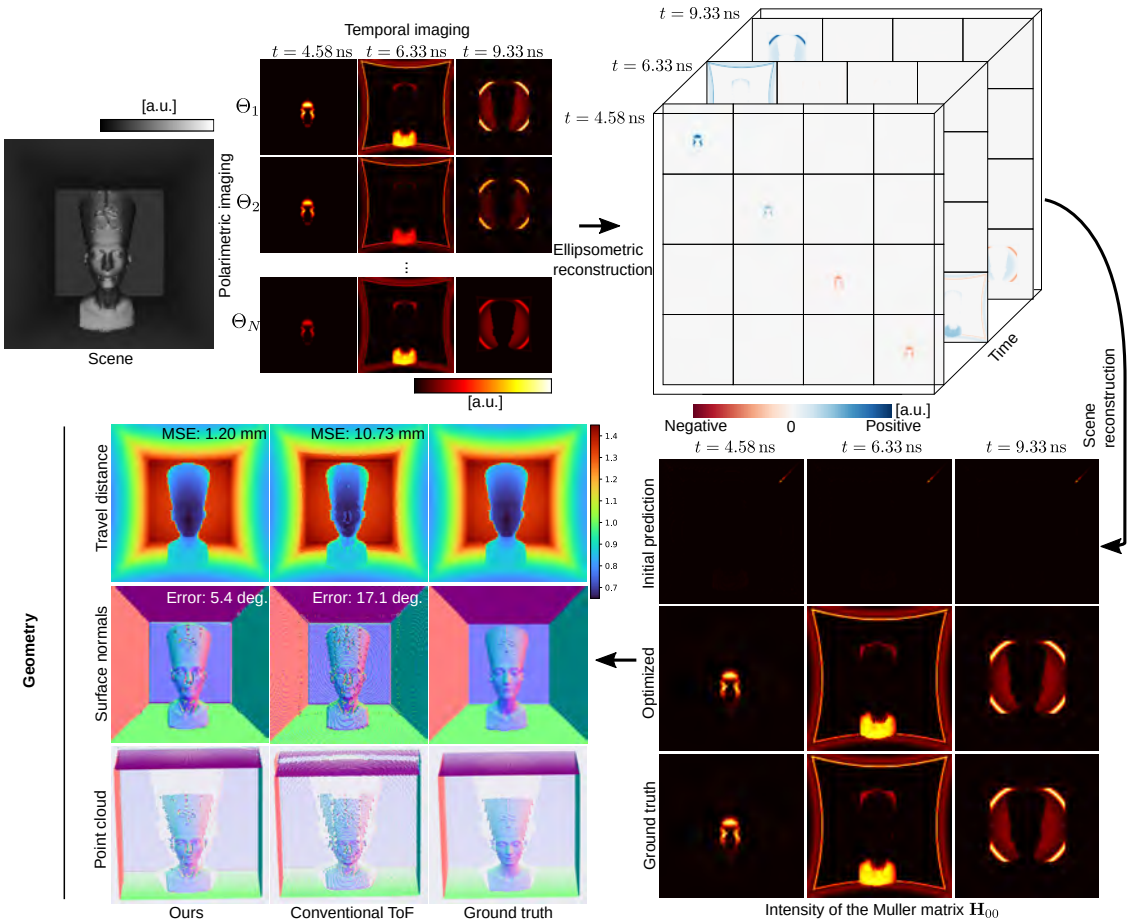
[4] Edward Collett. Field guide to polarization. Spie Bellingham, WA, 2005. 1, 2, 3

[5] Eric Compain, Stéphane Poirier, and Bernard Drevillon. General and self-consistent method for the calibration of polarization modulators, polarimeters, and mueller-matrix ellipsometers. *Applied optics*, 38(16):3490–3502, 1999. 3

[6] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 303–312, 1996. 4

[7] Eric Heitz. Understanding the masking-shadowing function in microfacet-based brdfs. *Journal of Computer Graphics Techniques*, 3(2):32–91, 2014. 2

[8] Anna Petrova-Mayor and Sarah Knudsen. Analysis and manipulation of the induced changes in the state of polarization by mirror scanners. *Applied Optics*, 56(15):4513–4521, 2017. 2



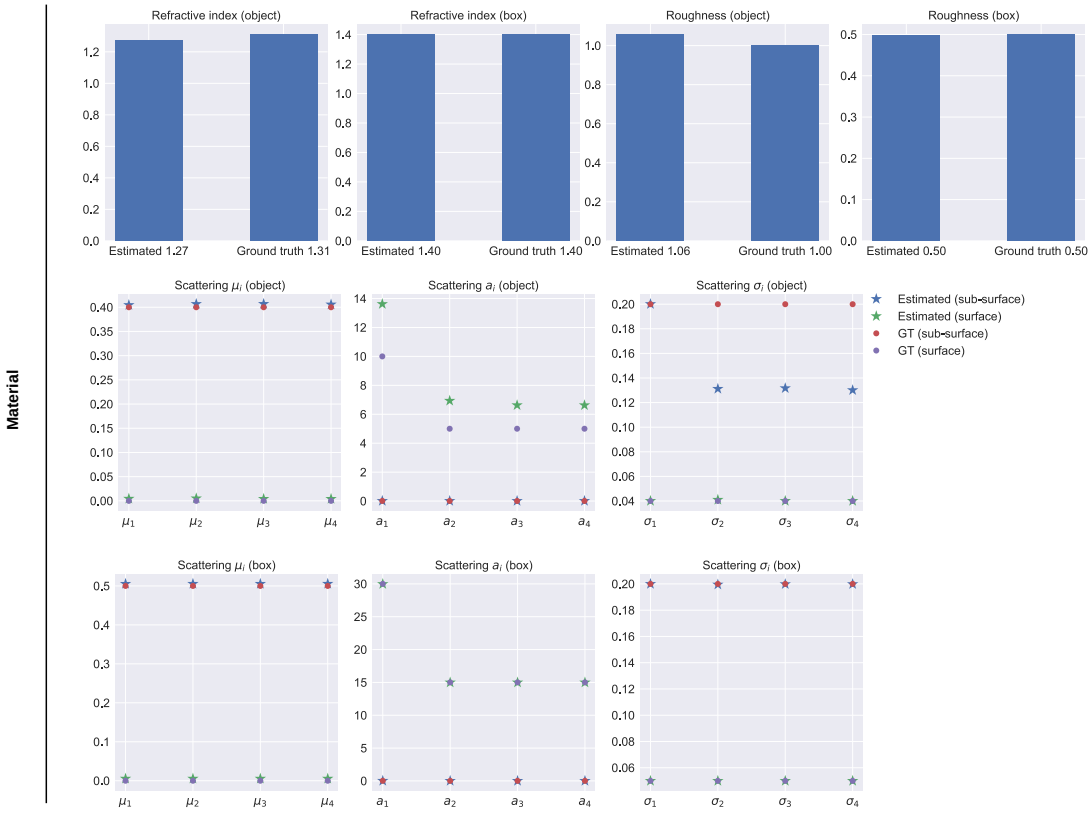
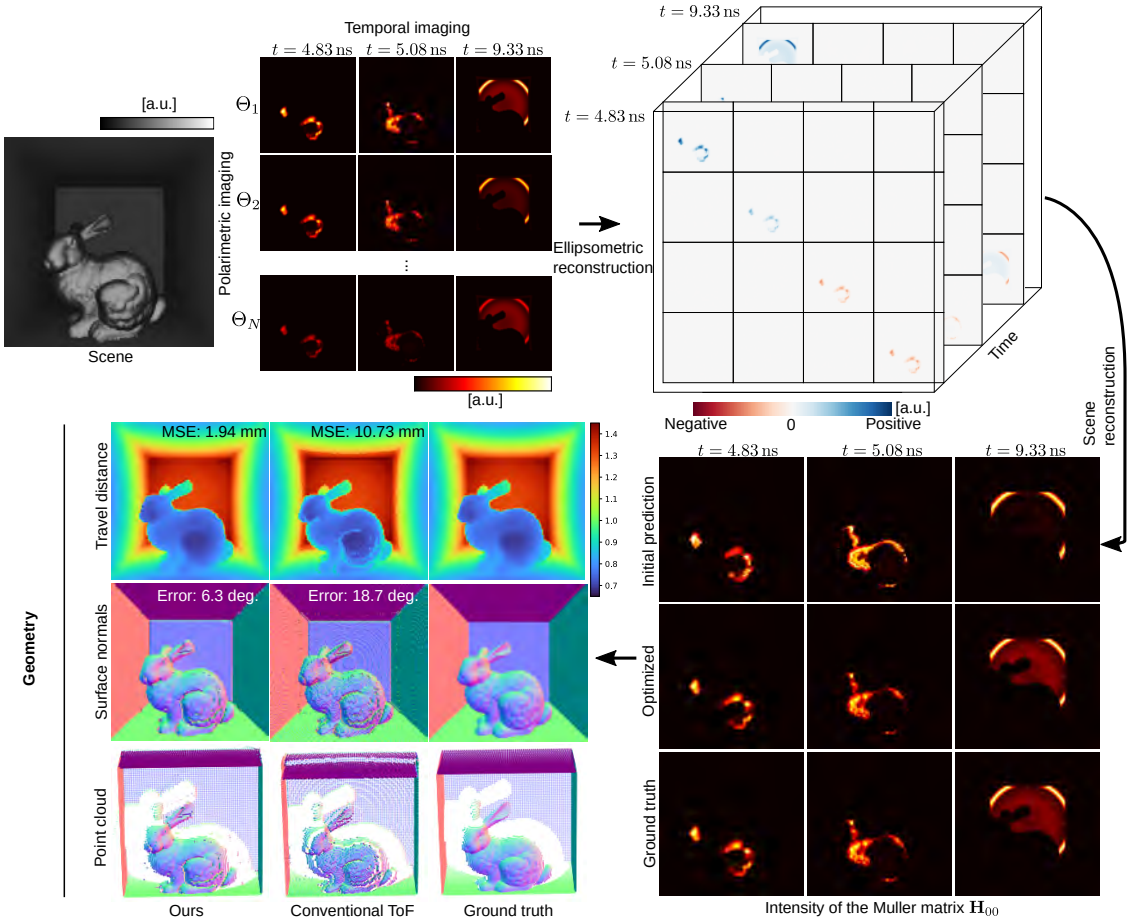


Figure 3. Reconstruction for the Synthetic Experiment on Scene (*Nefertiti*).



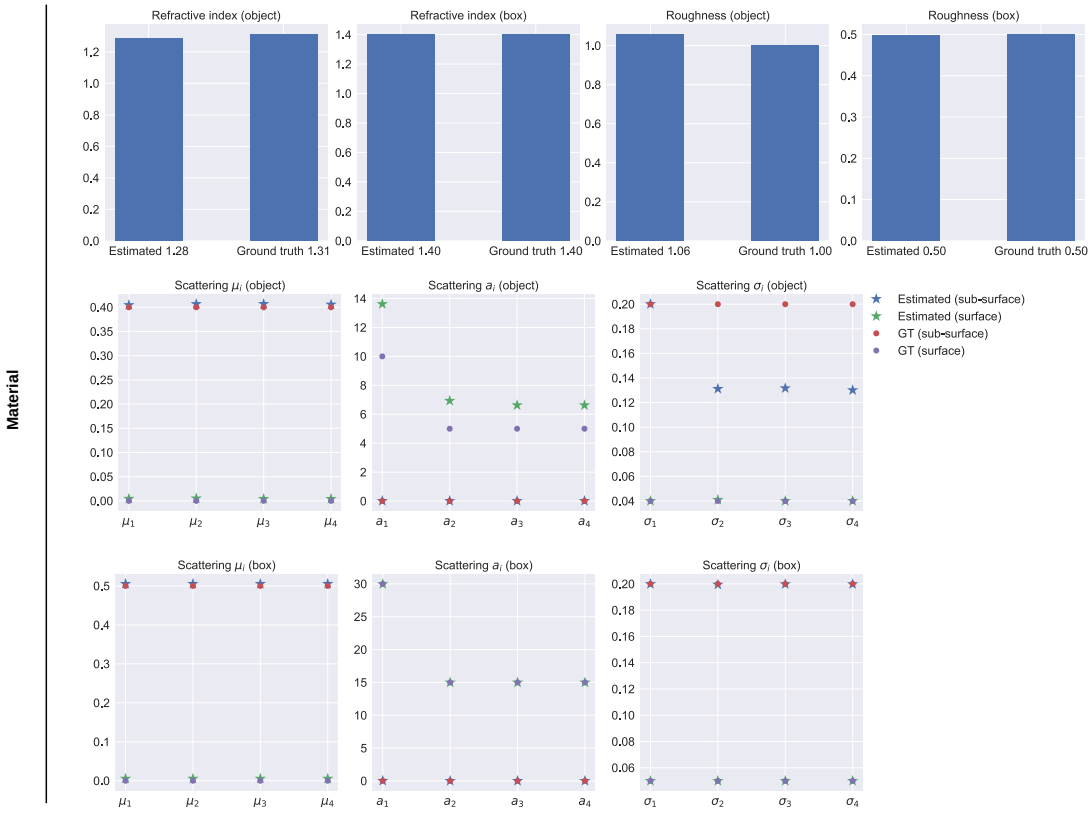
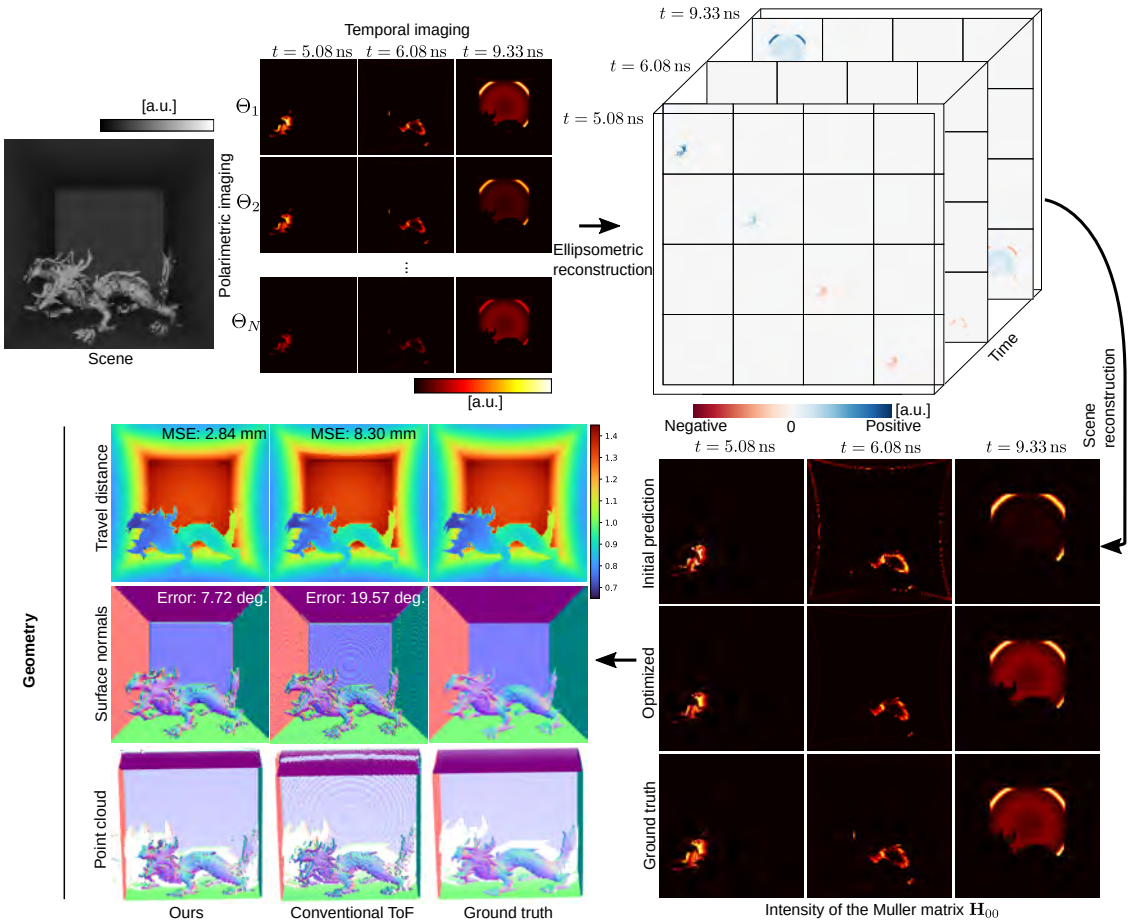


Figure 4. Reconstruction for the Synthetic Experiment on Scene (*Bunny*).



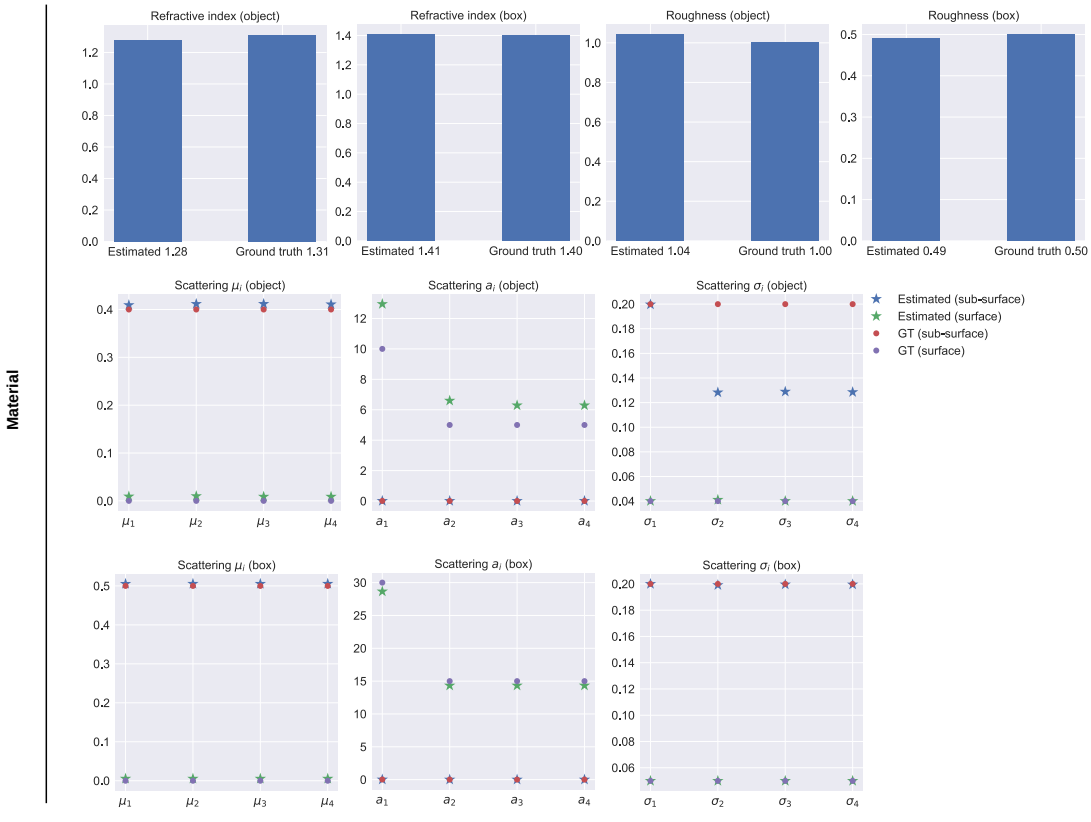


Figure 5. Reconstruction for the Synthetic Experiment on Scene (*Dragon*).

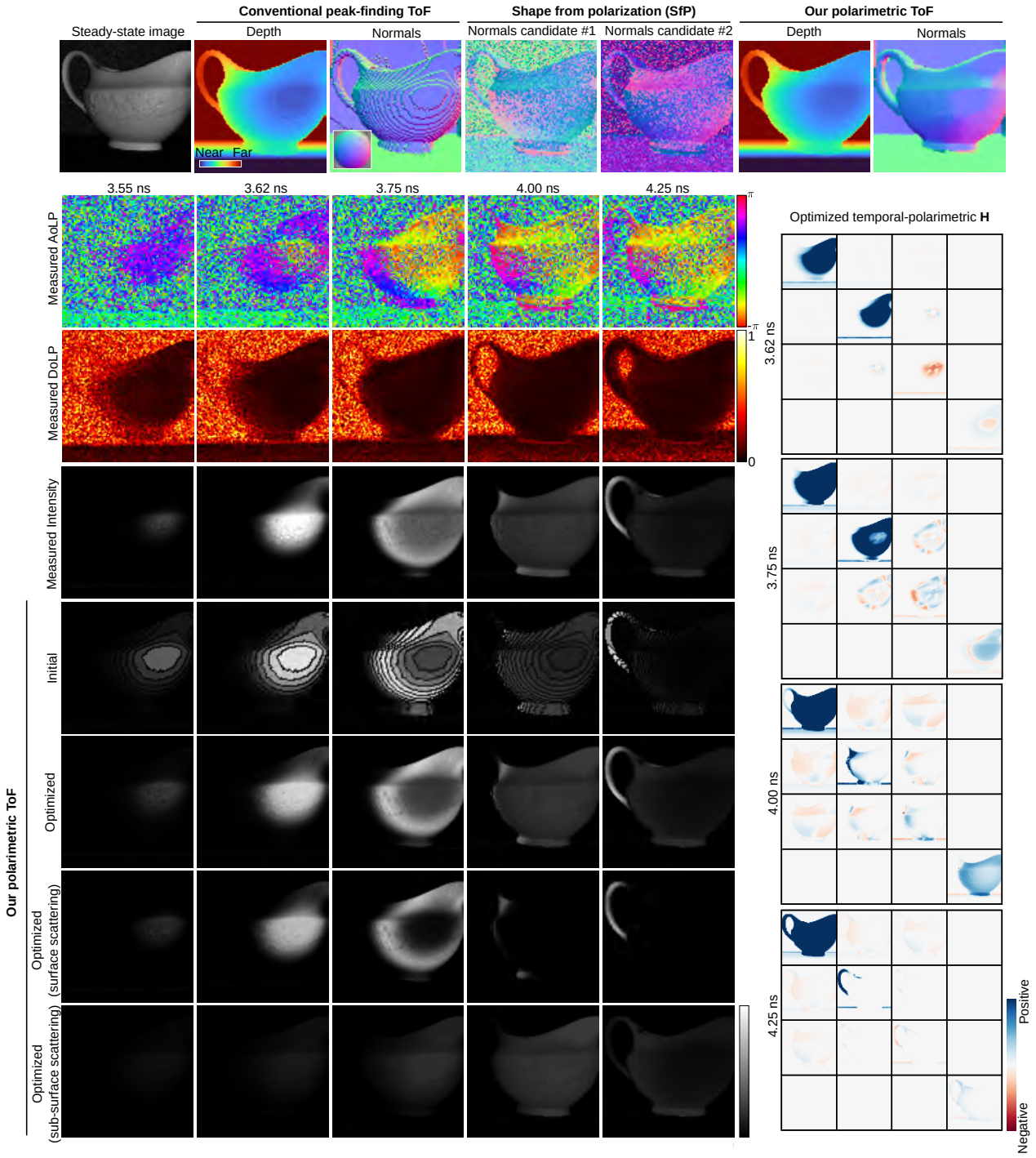


Figure 6. Scene Reconstruction for Experimental Measurement (*Pot*).

[9] Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance. Microfacet models for refraction through rough surfaces. *Rendering techniques*, 2007:18th, 2007. 2

[10] Di Wu, Andreas Velten, Matthew O'Toole, Belen Masia, Amit Agrawal, Qionghai Dai, and Ramesh Raskar. Decomposing global light transport using time of flight imaging. *Springer International Journal of Computer Vision (IJCV)*, 107(2):123–138, 2014. 5

[11] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000. 3

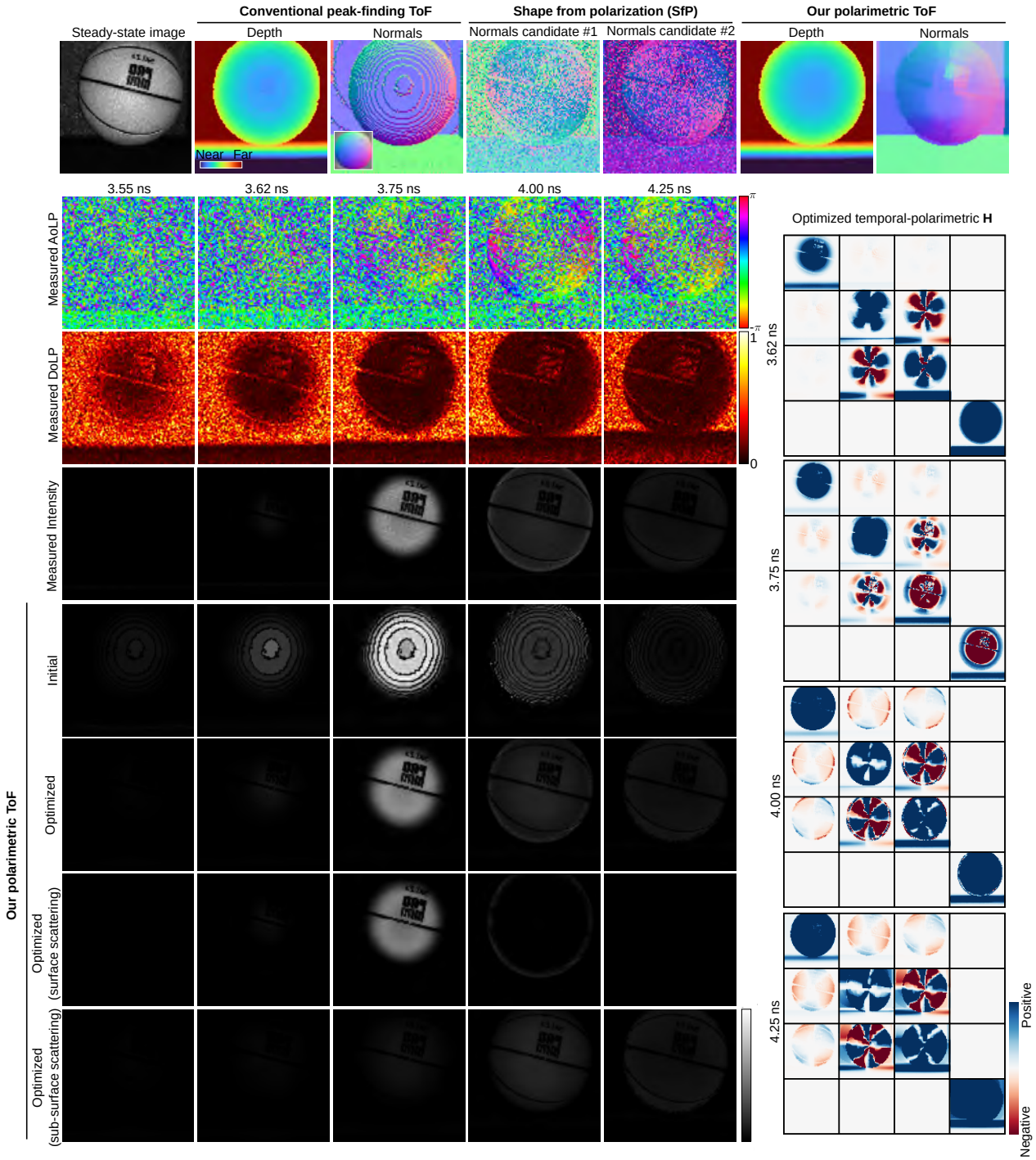


Figure 7. Scene Reconstruction for Experimental Measurement (*Basketball*).

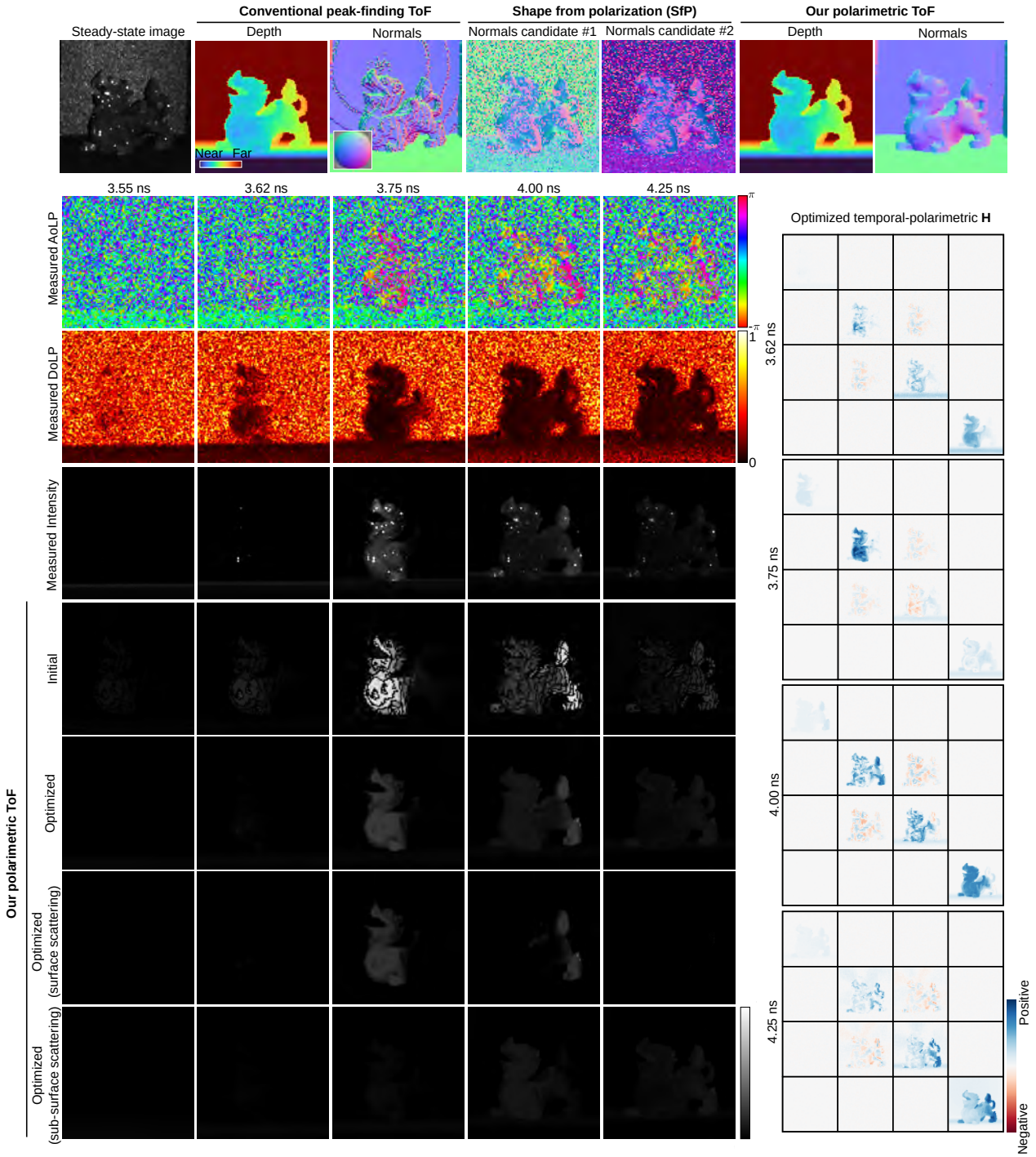


Figure 8. Scene Reconstruction for Experimental Measurement (*Dragon*).

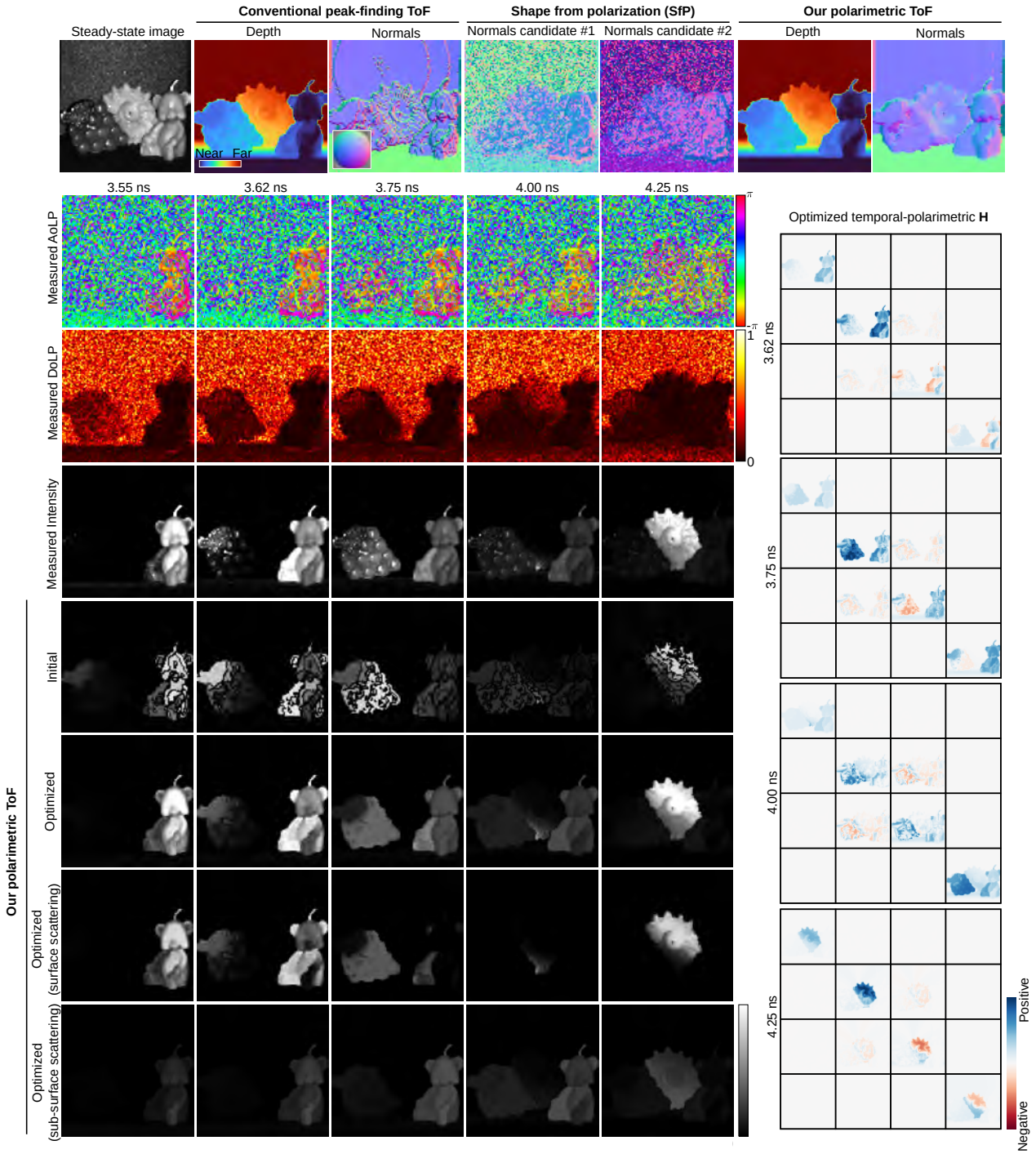


Figure 9. Scene Reconstruction for Experimental Measurement (*Candle*).